CHAPTER 1
INTRODUCTION

This chapter will give a detailed explanation about Automatic Number plate recognition system (ANPR).

1.1 Overview of Automatic Number Plate Recognition System

This section provides an introduction about the various terminologies in Automatic Number Plate Recognition (ANPR).

1.1.1 Images

An image is an artifact that depicts or records visual perception, for example a two-dimensional picture, that has a similar appearance to some subject–usually a physical object or a person, thus providing a depiction of it.

Images may be two-dimensional, such as a photograph, screen display, and as well as a three-dimensional, such as a statue or hologram. They may be captured by optical devices such as cameras, mirrors, lenses etc.

1.1.2 Image Processing

Image processing is referred to processing of a 2D picture by a computer. Basic definitions: An image defined in the “real world” is considered to be a function of two real variables, for example, a(x,y) with a as the amplitude (e.g. brightness) of the image at the real coordinate position (x,y).

Modern digital technology has made it possible to manipulate multi-dimensional signals with systems that range from simple digital circuits to advanced parallel computers. The goal of this manipulation can be divided into three categories:

Image Processing (image in -> image out)

Image Analysis (image in -> measurements out)

Image Understanding (image in -> high-level description out)

An image may be considered to contain sub-images sometimes referred to as regions-of-interest, ROIs, or simply regions. This concept reflects the fact that images frequently contain collections of objects each of which can be the basis for a region. In a sophisticated image processing system it should be possible to apply specific image processing operations to selected regions. Thus one part of an image might be processed to suppress motion blur while another part might be processed to improve color rendition.

1.1.3 Types of Image Processing

There are two types of methods used for Image Processing are Analog and Digital Image Processing.

Analog or visual techniques of image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. The image processing is not just confined to area that has to be studied but on knowledge of analyst. Association is another important tool in image processing through visual techniques. So analysts apply a combination of personal knowledge and collateral data to image processing.

Digital Processing techniques help in manipulation of the digital images by using computers. As raw data from imaging sensors from satellite platform contains deficiencies. To get over such flaws and to get originality of information, it has to undergo various phases of processing. The three general phases that all types of data have to undergo while using digital technique are Pre- processing, enhancement and display, information extraction.

1.1.4 Digital Image Processing Techniques

Digital image processing techniques are typically classified into three categories. These categories include image generation, enhancement, and restoration. Generation techniques help project and recognize a scanned image, while the process of enhancing an image involves improving contrast, brightness and hue. Restoration techniques help eliminate and correct errors that do not accurately reflect the original picture.

Many scanners that transmit images into computer programs use optical character recognition (OCR) technology. This converts the original image into text that the computer program will recognize. One of the main problems with any type of image processing technology is that some degree of editing is needed. Simple programs allow users to manipulate the image by cropping off unneeded space, changing the tint, brightness, and layout orientation.

Image generation is one of the digital image processing techniques that involves converting an image into some sort of ordered layout. For example, a scanner may pick up an image by creating a reflection. Digitization takes that reflection and attempts to arrange it into a series of pixels. Some forms of digitization make this arrangement by looking for variations in the amount of light recorded from the original image.

Enhancement is a broad category of digital image processing techniques that manipulate a digitized image. Computer programs may allow users to duplicate and change scanned images. For example, a color photograph can be converted to black and white. Likewise, parts of an image can be sliced and transposed into another.

1.1.5 Optical Character Recognition

Optical character recognition, usually abbreviated to OCR, is the mechanical or electronic conversion of scanned images of handwritten, typewritten or printed text into machine-encoded text. It is widely used as a form of data entry from some sort of original paper data source, whether documents, sales receipts, mail, or any number of printed records. It is a common method of digitizing printed texts so that they can be electronically searched, stored more compactly, displayed on-line, and used in machine processes such as machine translation, text-to-speech and text mining.

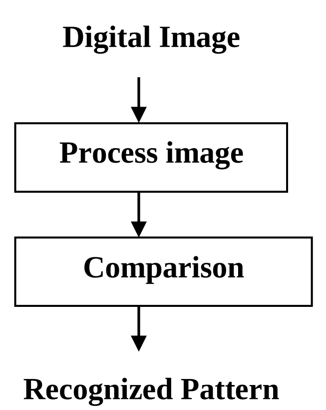


Figure 1.1 Block Diagram of Image Processing

1.2 PROBLEM STATEMENT

The task of the system is to recognise the numbers present in the number plate region of a vehicle in an embedded processing environment and give the same in a text format.

- **Input**

  The input of the process is an image that has the car and the number plate region present in it.

- **Output**

  Recognized numbers present in the number plate.

- **Objective**

  To provide an embedded solution for automatic number plate recognition in Indian conditions.

1.2.1 SCOPE

- The system obtains an image from the user and identifies the area in which the number plate is present by means of Double Phase Statistical Image Analysis. This algorithm involves two phases namely,
  - Band Clipping
  - Plate Clipping.

  The output of Band Clipping becomes the input of Plate Clipping algorithm.

- A given plate may have an inclination. We check if the pixels of the plate are in a straight line. If we have an inclination we apply De-skewing that straightens the inclined plate through rotations and shearing.
- The next phase is to segment the License plate. This is done with the help of Horizontal Projection.
- On the segmented plated we apply Piece Extraction Algorithm to obtain the character pixels.
- To classify the patterns obtained into their respective classes we make use of Neural Networks. We have two sets of elements, A set of possible descriptors and a set of possible classes.

1.2.2 LIMITATIONS

- Criminals can overcome ANPR by either making a copy of an identical vehicle number plate or changing the plate with a vehicle that is not reported stolen.
- Complex installation and start-up.
- High cost .

1.3 APPLICATIONS OF ANPR

Automatic Number Plate Recognition has a wide range of applications since the license number is the primary, most widely accepted, human readable, mandatory identifier of motor vehicles.

- ANPR provides automated access of the content of the number plate for computer systems managing databases and processing information of vehicle movements.

- Below we indicated some of the major applications, without the demand of completeness.

Parking

- One of the main applications of ANPR is parking automation and parking security: ticketless parking fee management, parking access automation, vehicle location guidance, car theft prevention, "lost ticket" fraud, fraud by changing tickets, simplified, partially or fully automated payment process, amongst many others.

Access Control

- Access control in general is a mechanisms for limiting access to areas and resources based on users' identities and their membership in various predefined groups. Access to limited zones, however, may also be managed based on the accessing vehicles alone, or together with personal identity. License plate recognition brings automation of vehicle access control management, providing increased security, car pool management for logistics, security guide assistance, event logging, event management, keeping access diary, possibilities for analysis and data mining.

Motorway Road Tolling

- Road Tolling means, that motorists pay directly for the usage of particular segment of road infrastructures. Tolls are a common way of funding the improvements of highways, motorways, roads and bridges: tolls are fees for services. Efficient road tolling increases the level of related road services by reducing travel time overhead, congestion and improve roadways quality. Also, efficient road tolling reduces fraud related to non-payment, makes charging effective, reduces required manpower to process events of exceptions. License plate recognition is mostly used as a very efficient enforcement tool, while there are road tolling systems based solely on license plate recognition too.

Border Control

- Border Control is an established state-coordinated effort to achieve operational control of the country's state border with the priority mission of supporting the homeland's security agains terrorism, illegal cross border traffic, smuggling and criminal activities. Efficient border control significantly decreases the rate of violent crime and increases the society's security. Automatic number plate recognition adds significant value by event logging, establishing investigate-able databases of border crossings, alarming on suspecious passings, at many more.

Journey Time Measurement

- Journey Time Measurement is a very efficient and widely usable method of understanding traffic, detecting conspicuous situations and events, etc. A computer vision based system has its well knwon downfalls in Journey Time Measurement, while Automatic Number Plate Recognition has provied its viability: vehicle journey times can be measured reliably by automatic number plate recognition-based systems. Data collected by license plate recognition systems can be used in many ways after processing: feeding back information to road users to increase traffic security, helping efficient law enforcement, optimising traffic routes, reducing costs and time, etc.

Law Enforcement

- Automatic number plate recognition is an ideal technology to be used for law enforcement purposes. It is able to automatically identify stolen cars based on the up-to date blacklist. Other very common law enforcement applications are red-light enforcement and overspeed charging and bus lane control.

1.4 ORGANIZATION OF THE REPORT

This section describes about various chapters included in this report.

Chapter 2: Literature review

This chapter presents some background information on the various areas involved in this research. These include various image processing paradigms.

Chapter 3: Proposed methodology

This chapter gives the complete architecture of the system. It explains the detailed design of the various phases and different modules used in the system. This section also describes about the hardware and software used to develop the number plate recognition system.

Chapter 4: Evaluation and results

This chapter explains about the results of various phases involved in the development of the system. It describes about the intermediate results that is used to test the performance of the system. Performance of the system in different environments is explained in this section.

Chapter 5: Conclusions and future directions

This chapter summarizes the main components of this system as well as the major conclusions gathered from the results of the evaluation. This section also covers how the system can be further improved.

CHAPTER 2
LITERATURE REVIEW

This section presents various image processing methods that can be used in ANPR systems in order to achieve the target of recognizing the numbers on the number plate.

2.1 CHARACTER RECOGNITION

The advancements in pattern recognition has accelerated recently due to the many emerging applications which are not only challenging, but also computationally more demanding, such evident in Optical Character Recognition (OCR), Document Classification, Computer Vision, Data Mining, Shape Recognition, and Biometric Authentication, for instance. The area of OCR is becoming an integral part of document scanners, and is used in many applications such as postal processing, script recognition, banking, security (i.e. passport authentication) and language identification. The research in this area has been ongoing for over half a century and the outcomes have been astounding with successful recognition rates for printed characters exceeding 99%, with significant improvements in performance for handwritten cursive character recognition where recognition rates have exceeded the 90% mark. Nowadays, many organizations are depending on OCR systems to eliminate the human interactions for better performance and efficiency.

The field of pattern recognition is a multidisciplinary field which forms the foundation of other fields, as for instance, Image Processing, Machine Vision, and Artificial Intelligence. Therefore, OCR cannot be applied without the help of Image Processing and/or Artificial Intelligence. Any OCR system goes through numerous phases including: data acquisition, preprocessing, feature extraction, classification and post-processing where the most crucial aspect is the preprocessing which is necessary to modify the data either to correct deficiencies in the data acquisition process due to limitations of the capturing device sensor, or to prepare the data for subsequent activities later in the description or classification stage.

Data preprocessing describes any type of processing performed on raw data to prepare it for another processing procedure. Hence, preprocessing is the preliminary step which transforms the data into a format that will be more easily and effectively processed. Therefore, the main task in preprocessing the captured data is to decrease the variation that causes a reduction in the recognition rate and increases the complexities, as for example, preprocessing of the input raw stroke of characters is crucial for the success of efficient character recognition systems. Thus, preprocessing is an essential stage prior to feature extraction since it controls the suitability of the results for the successive stages. The stages in a pattern recognition system are in a pipeline fashion meaning that each stage depends on the success of the previous stage in order to produce optimal/valid results. However, it is evident that the most appropriate feature vectors for the classification stage will only be produced with the facilitation from the preprocessing stage. The main objective of the preprocessing stage is to normalize and remove variations that would otherwise complicate the classification and reduce the recognition rate.

IMPORTANCE OF PREPROCESSING IN CHARACTER RECOGNITION:

The importance of the preprocessing stage of a character recognition system lies in its ability to remedy some of the problems that may occur due to some of the factors. Thus, the use of preprocessing techniques may enhance a document image preparing it for the next stage in a character recognition system. In order to achieve higher recognition rates, it is essential to have an effective preprocessing stage, therefore; using effective preprocessing algorithms makes the OCR system more robust mainly through accurate

2.2 PREPROCESSING TECHNIQUES

Preprocessing techniques are needed on colour, grey-level or binary document images containing text and/or graphics. In character recognition systems most of the applications use grey or binary images since processing colour images is computationally high. Such images may also contain non-uniform background and/or watermarks making it difficult to extract the document text from the image without performing some kind of preprocessing, therefore; the desired result from preprocessing is a binary image containing text only. Thus, to achieve this, several steps are needed, first, some image enhancement techniques to remove noise or correct the contrast in the image, second, thresholding to remove the background containing any scenes, watermarks and/or noise, third, page segmentation to separate graphics from text, fourth, character segmentation to separate characters from each other and, finally, morphological processing to enhance the characters in cases where thresholding and/or other preprocessing techniques eroded parts of the characters or added pixels to them. The above techniques present few of those which may be used in character recognition systems and in some applications; few or some of these techniques or others may be used at different stages of the OCR system. The rest of the chapter will present some of the techniques used during the preprocessing stage of a character recognition system.

2.2.1 IMAGE ENHANCEMENT TECHNIQUES

Image enhancement improves the quality of images for human perception by removing noise, reducing blurring, increasing contrast and providing more detail. This section will provide some of the techniques used in image enhancement.

2.2.1.1 SPATIAL IMAGE FILTERING OPERATIONS

In image processing, filters are mainly used to suppress either the high frequencies in the image, i.e. smoothing the image, or the low frequencies, i.e. enhancing or detecting edges in the image. Image restoration and enhancement techniques are described in both the spatial domain and frequency domain, i.e. Fourier transforms. However, Fourier transforms require substantial computations, and in some cases are not worth the effort. Multiplication in the frequency domain corresponds to convolution in the time and the spatial domain. Using a small convolution mask, such as 3x3, and convolving this mask over an image is much easier and faster than performing Fourier transforms and multiplication; therefore, only spatial filtering techniques will be presented in this chapter. Images captured often may be influenced by noise; however, the resulting images may not provide desired images for analysis. In addition, in images with acceptable quality, certain regions may need to be emphasized or highlighted. Spatial processing is classified into point processing and mask processing. Point processing involves the transformation of individual pixels independently of other pixels in the image. These simple operations are typically used to correct for defects in image acquisition hardware, for example to compensate for under/over exposed images. On the other hand, in mask processing, the pixel with its neighbourhood of pixels in a square or circle mask are involved in generating the pixel at (x, y) coordinates in the enhanced image.

2.2.1.2 POINT PROCESSING

Point processing modifies the values of the pixels in the original image to create the values of the corresponding pixels in the enhanced image this is expressed in equation (1).

$$O(x,y) = T[I(x,y)] \tag{1}$$

Where, I(x, y) is the original (input) image, O(x, y) is the enhanced image and T describes the transformation between the two images. Some of the point processing techniques include: contrast stretching, global thresholding, histogram equalisation, log transformations and power law transformations. Some mask processing techniques include averaging filters, sharpening filters, local thresholding… etc.

2.2.1.3 GLOBAL IMAGE THRESHOLDING

Image thresholding is the process of separating the information (objects) of an image from its background, hence, thresholding is usually applied to grey-level or colour document scanned images. Thresholding can be categorised into two main categories: global and local. Global thresholding methods choose one threshold value for the entire document image, which is often based on the estimation of the background level from the intensity histogram of the image; hence, it is considered a point processing operation. On the other hand, local adaptive thresholding uses different values for each pixel according to the local area information. There are hundreds of thresholding algorithms which have been published in the literature and presenting all methods would need several books, therefore, the purpose here is to present some of the well-known methods.

Global thresholding methods are used to automatically reduce a grey-level image to a binary image. The images applied to such methods are assumed to have two classes of pixels (foreground and background). The purpose of a global thresholding method is to automatically specify a threshold value, T, where the pixel values below it are considered foreground and the values above are background. A simple method would be to choose the mean or median value of all the pixels in the input image, the mean or median will work well as the threshold, however, this will generally not be the case especially if the pixels are not uniformly distributed in an image. A more sophisticated approach might be to create a histogram of the image pixel intensities and use the valley point (minimum) as the threshold. The histogram approach assumes that there is some average value for the background and object pixels, but that the actual pixel values have some variation around these average values. However, this may be computationally expensive, and image histograms may not have clearly defined valley points, often making the selection of an accurate threshold difficult. One method that is relatively simple and does not require much specific knowledge of the image is the iterative method (Gonzalez, et al., 2004) which is explained below.

The iterative procedure is

> Step 1: Select an initial threshold value (T), randomly or according to any other method desired such as the mean or median value of the pixels in the image.

Step 2: Segment the image, using T, into object and background pixels. R1 (background region) consists of pixels with intensity values $\geq$ T and R2 (objects region) consists of pixels with intensity $<$ T.

Step 3: Calculate the average of each region, $\mu 1$ and $\mu 2$ for regions R1 and R2, respectively.

Step 4: Compute the new threshold value T as given in equation (2).

$$T=1/2(\mu 1 + \mu 2) \qquad (2)$$

Step 5: Repeat the steps from 2 – 4 using the new T until the new threshold matches the one before it.

In the literature, many thresholding methods have been published, for example, Sahoo et al. compared the performance of more than 20 global thresholding algorithms using uniformly or shape measures. The comparison showed that Otsu class separability method gave best performance (Sahoo et al., 1988; Otsu, 1979). On the other hand, in an evaluation for change detection by Rosin & Ioannidis concluded that the Otsu algorithm performed very poorly compared to other global methods (Rosin & Ioannidis, 2003, Otsu, 1979). The OCR goal-directed evaluation study by Trier and Jain examined four global techniques showing that the Otsu method outperformed the other methods investigated in the study (Trier & Jain, 1995). In addition, Fischer compared 15 global methods and confirmed that the Otsu method is preferred in document image processing (Fischer, 2000). The Otsu method is one of the widely used techniques used to convert a grey-level image into a binary image then calculates the optimum threshold separating those two classes so that their combined spread (intra-class variance) is minimal.

2.2.1.4 Histogram processing

Histogram processing is used in image enhancement and can be useful in image compression and segmentation processing. A histogram simply plots the frequency at which each grey-level occurs from 0 (black) to 255 (white). Scanned or captured images may have a limited range of colours, or are lacking contrast (details). Enhancing the image by histogram processing can allow for improved detail, but can also aid other machine vision operations, such as segmentation. Thus, histogram processing should be the initial step in preprocessing. Histogram equalisation and histogram specification (matching) are two methods widely used to modify the histogram of an image to produce a much better image.

Histogram equalisation

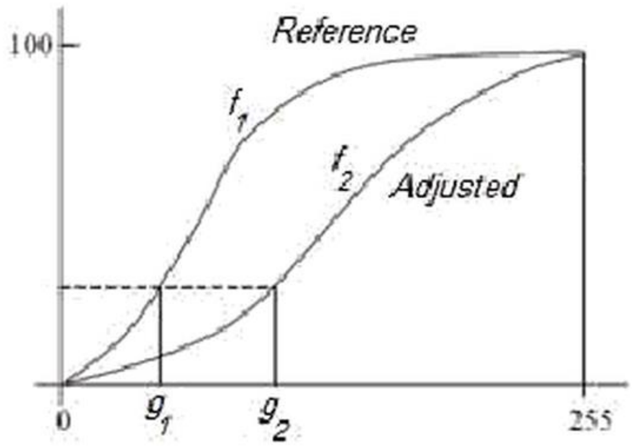
Histogram equalisation is considered a global technique. It stretches the histogram across the entire spectrum of pixels (0 – 255). It increases the contrast of images for the finality of human inspection and can be applied to normalize illumination variations in image understanding problems. This process is quite simple and for each brightness level j in the original image, the new pixel level value (k) is calculated as given in equation (3).

$$K = \Sigma\, Ni / T \tag{3}$$

where the sum counts the number of pixels in the image (by integrating the histogram) with brightness equal to or less than j, and T is the total number of pixels (Russ, 2007). In addition, histogram equalisation is one of the operations that can be applied to obtain new images based on histogram specification or modification.

Histogram specification (Matching)

Histogram matching is a method in image processing of colour adjustment of two images using their image histograms.



Histogram modification is the matching of the cumulative function f2 of the image to be adjusted to the Cumulative Distribution Function (CDF) of the reference image f1. Histogram modification is done by first computing the histograms of both images then the CDFs of both the reference (f1) and to be adjusted (f2) images are calculated. This output of the histogram matching is obtained by matching the closest CDF f2 to the reference image CDF f1. Then for each grey-level g1 the grey-level g2 is calculated for which f1 (g1) = f2 (g2) as shown in above Figure, and this is the result of histogram matching function M(g1) = g2 (Horn & Woodham, 1979).

2.2.2 Smoothing (Low-pass) filters

Average or mean filter is a simple, intuitive and easy to implement method of smoothing images, i.e. reducing the amount of intensity variation between one pixel and the next. It is often used to reduce noise in images. In general, the mean filter acts as a low-pass frequency filter and, therefore, reduces the spatial intensity derivatives present in the image. The idea of mean filtering is simply to replace each pixel value in an image with the mean (`average') value of its neighbours, including itself. This has the effect of eliminating pixel values which are unrepresentative of their surroundings. Mean filtering is usually thought of as a convolution filter. Like other convolutions it is based around a kernel, which represents the shape and size of the neighbourhood to be sampled when calculating the mean. Often a 3×3 square kernel/mask is used, as shown in Fig. 3, although larger masks can be used (e.g. 5×5, 7x7, 9x9 ...) for more severe smoothing. Note that, a small kernel can be applied more than once in order to produce a similar, but not identical, effect as a single pass with a larger kernel. Also, the elements of the mask must be positive and hence the size of the mask determines the degree of smoothing. Therefore, the larger the window size used a blurring effect is produced causing small objects to merge with the background of the image (Nixon&Aguado, 2008).

1/9 x

1	1	1
1	1	1
1	1	1

The center coefficient of the mask is very important and other pixels are inversely weighted as a function of their distance from the center of the mask. The basic strategy behind weighting the center point the highest and then reducing the value of the coefficients as a function of increasing distance from the origin is simply an attempt to reduce blurring in the smoothing process.

2.2.3 Sharpening (High-pass) filter

A sharpening filter is used to emphasize the fine details of an image (i.e., provides the opposite effect of smoothing). The points of high contrast can be detected by computing intensity differences in local image regions.

The weights of the mask are both positive and negative. When the mask is over an area of constant or slowly varying grey-level, the result of convolution will be close to zero. When grey-level is varying rapidly within the neighbourhood, the result of convolution will be a large number. Typically, such points form the border between different objects or scene parts (i.e. edge). An example of a sharpening filter is the Laplacian filter which is defined in equation (4) below.

$$\nabla^2 f = [f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1)] - 4f(x,y)$$

This implementation can be applied at all points (x,y) in an image by convolving the image with the following spatial mask Fig. 4(a) with an alternative definition of the digital second derivatives which takes into account the diagonal elements and can be implemented by the mask in the Figure.

0	1	0
1	-4	1
0	1	0

(a)

1	1	1
1	-8	1
1	1	1

(b)

The Laplacian filter is a derivative operator which sharpens the image, but drives constant areas to zero; therefore, adding the original image back restores the grey-level tonality, equation (5).

$$g(x,y) = f(x,y) + c[\nabla^2 f(x,y)]$$

Where, f(x,y) is the input image, g(x,y) is the output image and c is 1 if the centre coefficient of the mask is positive, or -1 if it is negative (Gonzales and Woods, 2002).

2.3 LOG TRANSFORMATIONS

The general form of the log transformation is equation (6).

$$s = c \log (1 + r) \tag{6}$$

Where c is a constant and it is assumed that $r \geq 0$. This transformation maps a narrow range of low grey-level values in the input image into a wider range of output levels and vice versa (Gonzalez et al., 2004).

2.4 POWER LAW TRANSFORMATION

Power-law transformations have the general form shown in equation (7).

$$S = C (r + \varepsilon) ^\wedge \gamma \tag{7}$$

where c and γ are positive constants and □ is an offset which is usually ignored since it is due to display calibration. Therefore; $S = C . r ^\wedge \gamma$, where values of $0 < \gamma < 1$ map a narrow range of dark input values into a wider range of output values, with the opposite being true for values of γ greater than 1. This shows that the power-law transformations are much more versatile in such application than the log transformation. However, the log function has the important characteristic that it compresses the dynamic range of images with large variations in pixel values. Due to the variety of devices used for image capture, printing, and display respond according to the power law exponent, gamma, (γ), this factor needs to be corrected, thus power-law response phenomena or gamma correction which is given by

$S = C . r ^\wedge 1/\gamma$ (Gonzalez et al., 2004).

2.5 MASK PROCESSING

In mask processing, a pixel value is computed from the pixel value in the original image and the values of pixels in its vicinity. It is a more costly operation than simple point processing, but more powerful. The application of a mask to an input image produces an output image of the same size as the input.

2.6 NOISE REMOVAL

The advancements in technology produced image acquisition devices with better improvements. While modern technology has made it possible to reduce the noise levels associated with various electro-optical devices to almost negligible levels, there are still some noise sources which cannot be eliminated. Images acquired through modern sensors may be contaminated by a variety of noise sources. By noise we refer to stochastic variations as opposed to deterministic distortions, such as shading or lack of focus. There are different types of noise that are related to the electronic capturing devices or the light source used such types of noise are photon, thermal, On-Chip electronic and quantisation. Most of the noise may be eliminated by the capturing sensors or the CCD cameras. Document analysis systems benefit from the reduction of noise in the preprocessing stage this can provide a substantial improvement in the reliability and robustness of the feature extraction and recognition stages of the OCR system. A common manifestation of noise in binary images takes the form of isolated pixels, salt-and-pepper noise or speckle noise, thus; the processing of removing this type of noise is called filling, where each isolated pixel saltand-pepper “island” is filled in by the surrounding “sea” (O’Gorman, et al., 2008). In greylevel images or median filters and low-pass filters such as average or Gaussian blur filters proved to eliminate isolated pixel noise. Gaussian blur and average filters are a better choice to provide smooth texture to the image. On the other hand, periodic noise which manifests itself as impulse-like bursts which often are visible in the Fourier spectrum can be filtered using notch filtering. The transfer function of a Butterworth notch filter of order n, H(u,v),is given by equation.

$$H(u,v) = \frac{1}{1 + \left[\frac{D_0^2}{D_1(u,v) D_2(u,v)} \right]^n}$$

Where,

$$D_1(u,v) = [(u - M/2 - u_0)^2 + (v - N/2 - v_0)^2]^{1/2}$$

And

$$D_2(u,v) = [(u - M/2 + u_0)^2 + (v - N/2 + v_0)^2]^{1/2}$$

where (µ0, υ0) and by symmetry (-µ0, -υ0) are the locations of the notches and D is their radius, equations 14 - 15. The filter is specified with respect to the centre of the frequency rectangle. (Gonzalez et al., 2004).

2.7 SKEW DETECTION/CORRECTION

Due to the possibility of rotation of the input image and the sensitivity of many document image analysis methods to rotation of the image, document skew should be corrected. Skew detection techniques can be roughly classified into the following groups: analysis of projection profile, Hough transform, connected components, clustering, and Correlation between lines techniques. The survey by Hull and Taylor, investigated twenty-five different methods for document image skew detection. The methods include approaches based on Hough Transform analysis, projection profile, feature point distribution and orientation sensitive feature analysis. The survey concluded that most of the techniques reported a range of up to 0.1 degrees accuracy, evidencing a strong need for further work in this area to help show the strengths and weaknesses of individual algorithms (Hull & Taylor, 1998). In addition, there are new techniques emerging for specific applications such as the method of Al-Shatnawi and Omar which is based on the center of gravity for dealing with Arabic document images (Al-Shatnawi & Omar, 2009). Therefore, the choice of using a skew detection/correction technique depends on the application and the type of images used.

2.8 GENERAL OBSERVATIONS

Suri, walia and verma in their paper, approached the use of sobel edge detection technique to detect the edges in the image [12].

Liu Ying presents a License Plate Recognition System which uses Adaptive Algorithm that can adapt in the event of changes in environment or circumstances [13].

Goshal proposes a number plate recognition system using segmentation techniques for the separation of the characters that is used in different countries. It also approached syntactical method for different countries to identify and recognize the number plates [14].

Lan Wang uses corpus and data collection which contains the substitution, insertion and deletion of phones in order to produce the correct word in spite of it being mispronounced[15].

Thangarajan proposes a system in which the phones of the Tamil words spoken are identified using the acoustic notes and the phoneme set of Tamil which has 44 phones. He also uses the word corpuses for comparisons[17].

Saraswathi uses the rules of grammar in Tamil language to eliminate the possibility of substituting the letters which cannot occur in some location in a word. If it is not a valid phoneme then it implies an error in segmentation at that point[16].

CHAPTER 3

PROPOSED METHODOLOGY

This section describes about the system architecture and the different modules used in developing multilingual speech recognition system.

3.1 PROPOSED SYSTEM

In this project a car arrives at the point of detection. On arrival the first step is to capture the image of the car. The image is captured with the help of a web camera.

After capturing an image the first phase is to detect the presence of a plate. This is to be done by Double Phase Statistical Image Analysis. This algorithm involves two phases namely, Band Clipping and Plate Clipping. The output of Band Clipping becomes the input of Plate Clipping algorithm.

A given plate may have an inclination. We check if the pixels of the plate are in a straight line. If we have an inclination we apply De-skewing that straightens the inclined plate through rotations and shearing.

The next phase is to segment the Licence plate. This is done with the help of Horizontal Projection. On the segmented plated we apply Piece Extraction

Algorithm to obtain the character pixels (the numbers).

To classify the patterns obtained into their respective classes we make use of Neural Networks. We have two sets of elements, a set of possible descriptors and a set of possible classes. The function which is to be developed matches this pattern to the corresponding class effectively.

To achieve all these functionalities the web camera is connected to Wand board - Dual. The images of this camera is given as an input to the board which does the processing work and in case of a match in the database it gives an output to a stepper motor which opens the gate for entry. In case, we don't find a match we display an interface to add a new user.

3.2 HARDWARE & SOFTWARE USED

i. A system with Core 2 Duo Processor with 1GB RAM.

ii. Fedora [18] core 12.

iii. Approximately 17 MB of free disk space for tools

iv. Hidden markov model ToolKit (HTK)-Speech recognition toolkit

v. NetbeansIDE - 6.7.1 - used for developing GUI in Linux.

vi. Wavesurfer[20] - is an open source tool for sound visualization, transcription and other manipulations.

3.3 MODULES

The different modules used in developing the system is explained in Figure 3.3

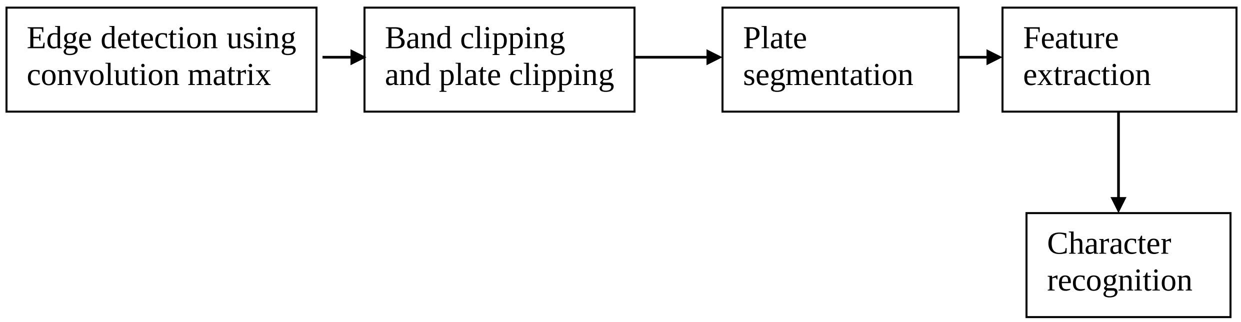


Figure 3.3 Module diagram

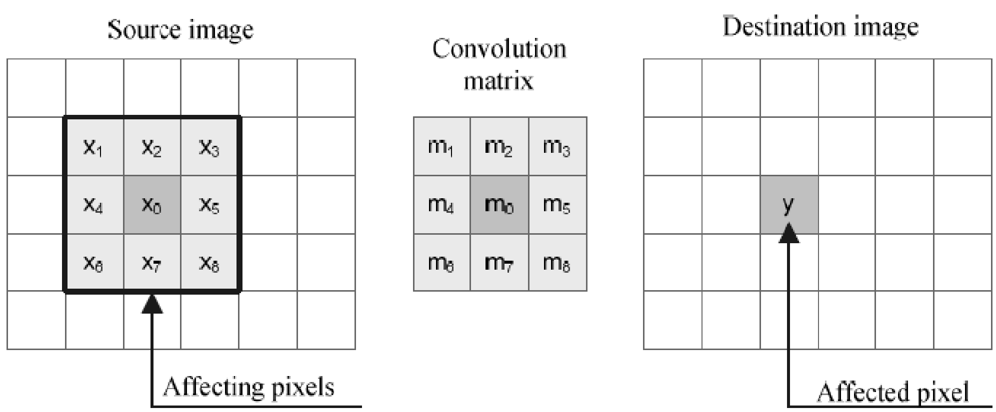
3.3.1Module 1 – Edge detection

The first step in a process of automatic number plate recognition is a detection of a number plate area. This problematic includes algorithms that are able to detect a rectangular area of the number plate in an original image. Humans define a number plate in a natural language as a "small plastic or metal plate attached to a vehicle for official identification purposes", but machines do not understand this definition as well as they do not understand what "vehicle", "road", or whatever else is. Because of this, there is a need to find an alternative definition of a number plate based on descriptors that will be comprehensible for machines. Let us define the number plate as a "rectangular area with increased occurrence of horizontal and vertical edges". The high density of horizontal and vertical edges on a small area is in many cases caused by contrast characters of a number plate, but not in every case. This process can sometimes detect a wrong area that does not correspond to a number plate. Because of this, we often detect several candidates for the plate by this algorithm, and then we choose the best one by a further heuristic analysis. The detection of a number plate area consists of a series of convolve operations. Modified snapshot is then projected into axes *x* and *y*. These projections are used to determine an area of a number plate. We can use a periodical convolution of the function *f* with specific types of matrices **g** to detect various types of edges in an image:

$$(g * f)(x, y) = \sum_{(a,b)\in A} g(a, b) f(x - a, y - b),$$

Where f is an image, and g some convolution kernel.

The convolution matrix defines how the specific pixel is affected by neighboring pixels in the process of convolution. Individual cells in the matrix represent the neighbors related to the pixel situated in the centre of the matrix. The pixel represented by the cell y in the destination image is affected by the pixels x0…x8 according to the formula:



3.3.1.1 Horizontal and vertical edge detection

To detect horizontal and vertical edges, we convolve source image with the following matrices. The convolution matrices are usually much smaller than the actual image. Also, we can use bigger matrices to detect rougher edges.

$$\begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

3.3.1.2 Sobel edge detector

The Sobel edge detector uses a pair of 3x3 convolution matrices. The first is dedicated for evaluation of vertical edges, and the second for evaluation of horizontal edges.

$$\begin{pmatrix} -1 & -2 & 1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \qquad \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$$

3.3.1.3 Horizontal and vertical image projection

After the series of convolution operations, we can detect an area of the number plate according to a statistics of the snapshot. There are various methods of statistical analysis. One of them is a horizontal and vertical projection of an image into the axes *x* and *y*. The vertical projection of the image is a graph, which represents an overall magnitude of the image according to the axis *y*. If we compute the vertical projection of the image after the application of the vertical edge detection filter, the magnitude of certain point represents the occurrence of vertical edges at that point. Then, the vertical projection of so transformed image can be used for a vertical localization of the number plate. The horizontal projection represents an overall magnitude of the image mapped to the axis *x*.

Let an input image be defined by a discrete function f (x, y). Then, a vertical projection p(y) of the function f at a point y is a summary of all pixel magnitudes in the y^{th} row of the input image. Similarly, a horizontal projection at a point x of that function is a summary of all magnitudes in the x^{th} column.

$$P(x) = \sum f(x, j)$$

3.3.2 Module 2 – Number plate detection

The detection of the number plate area consists of a "band clipping" and a "plate clipping". The band clipping is an operation, which is used to detect and clip the vertical area of the number plate (so-called band) by analysis of the vertical projection of the snapshot. The plate clipping is a consequent operation, which is used to detect and clip the plate from the band (not from the whole snapshot) by a horizontal analysis of such band.

3.3.2.1 Band clipping

The band clipping is a vertical selection of the snapshot according to the analysis of a graph of vertical projection. If h is the height of the analyzed image, the corresponding vertical projection $p(y)$ contains h values, such as y € (0,h-1)

3.3.2.2 Plate clipping

There is a strong analogy in a principle between the band and plate clipping. The plate clipping is based on a horizontal projection of band. At first, the band must be processed by a vertical detection filter. If w is a width of the band (or a width of the analyzed image), the corresponding horizontal projection $p(x)$ contains w values.

$$p(x) = \sum f(x, j)$$

$p(x)$ is a projection of the band, not of the whole image. This can be achieved by a summation in interval < Yb0 , Yb1 >, which represents the vertical boundaries of the band.

3.3.2.3 Deskewing

The captured rectangular plate can be rotated and skewed in many ways due to the positioning of vehicle towards the camera. Since the skew significantly degrades the recognition abilities, it is important to implement additional mechanisms, which are able to detect and correct skewed plates.

The fundamental problem of this mechanism is to determine an angle, under which the plate is skewed. Then, deskewing of so evaluated plate can be realized by a trivial affine transformation. It is important to understand the difference between the "sheared" and "rotated" rectangular plate. The number plate is an object in three-dimensional space, which is projected into the two dimensional snapshot during the capture. The positioning of the object can sometimes cause the skew of angles and proportions.

3.3.2.3.1 Detection of skew

Hough transform is a special operation, which is used to extract features of a specific shape within a picture. The classical Hough transform is used for the detection of lines. The Hough transform is widely used for miscellaneous purposes in the problematic of machine vision, but we have used it to detect the skew of captured plate, and also to compute an angle of skew. It is important to know, that Hough transform does not distinguish between the concepts such as "rotation" and "shear". The Hough transform can be used only to compute an approximate angle of image in a two-dimensional domain.

3.3.2.3.2 Correction of skew

As the skew detection based on Hough transform does not distinguish between the shear and rotation, it is important to choose the proper deskewing operation. In practice, plates are sheared in more cases than rotated. To correct the plate sheared by the angle Θ we use the affine transformation to shear it by the negative angle $-\Theta$.

For this transformation, we define a transformation matrix **A :**

$$\mathbf{A} = \begin{bmatrix} 1 & S_y & 0 \\ S_x & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\tan(\theta) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where S_x and S_y are shear factors. The S_x is always zero, because we shear the plate only in a direction of the Y-axis.

3.3.3 Module 3 – plate segmentation

The next step after the detection of the number plate area is a segmentation of the plate. The segmentation is one of the most important processes in the automatic number plate recognition, because all further steps rely on it. If the segmentation fails, a character can be improperly divided into two pieces, or two characters can be improperly merged together. We can use a horizontal projection of a number plate for the segmentation. If we assume only one-row plates, the segmentation is a process of finding horizontal boundaries between characters.

The second phase of the segmentation is an enhancement of segments. The segment of a plate contains besides the character also undesirable elements such as dots and stretches as well as redundant space on the sides of character. There is a need to eliminate these elements and extract only the character.

3.3.3.1 Segmentation of plate using a horizontal projection

We compute a horizontal projection $p(x)$ of the plate $f(x, y)$. We use this projection to determine horizontal boundaries between segmented characters. These boundaries correspond to peaks in the graph of the horizontal projection.

The algorithm of segmentation iteratively finds the maximum peak in the graph of vertical projection. The peak is treated as a space between characters, if it meets some additional conditions, such as height of peak.

The algorithm then zeroizes the peak and iteratively repeats this process until no further space is found.

3.3.3.2 Character extraction

The segment of plate contains besides the character also redundant space and other undesirable elements. Since the segment has been processed by an adaptive thresholding filter, it contains only black and white pixels. The neighboring pixels are grouped together into larger pieces, and one of them is a character. Our goal is to divide the segment into the several pieces, and keep only one piece representing the regular character. This has been done with the help of piece extraction algorithm.

Algorithm:

Let the segment be defined by a discrete function $f(x, y)$ in the relative coordinate system, such as [0,0] is an upper left corner of the segment, and $[w-1, h-1]$ is a bottom right corner, where w and h are dimensions of the segment. The value of $f(x, y)$ is "1" for the black pixels, and "0" for the white space.

The piece P is a set of all neighboring pixels $[x, y]$, which represents a continuous element of black pixels. The pixel $[x, y]$ belongs to the piece P if there is at least one pixel $[x', y']$ from the P , such as $[x, y]$ and $[x', y']$ are neighbors.

In practice, we use the number of pixels in a set P is considered. Pieces with a higher value of pixels will be preferred.

3.3.3.3 Character normalization

Segmented characters have very much variation in size. In this phase, all the characters are normalized to predefined height (Vertical Length) in pixel. As the characters always have variable width (Horizontal Length), each character image is normalized to a size of 5 X 13, by image mapping technique.

3.3.4 Module 4 – Feature extraction

To recognize a character from a bitmap representation, there is a need to extract feature descriptors of such bitmap. As an extraction method significantly affects the quality of whole OCR process, it is very important to extract features. The first step is a normalization of a brightness and contrast of processed image segments. The characters contained in the image segments must be then resized to uniform dimensions. After that, the feature extraction algorithm extracts appropriate descriptors from the normalized characters. This chapter deals with the methods used in the process of normalization.

3.3.4.1 Histogram normalization

The histogram normalization is a method used to re-distribute intensities on the histogram of the character segments. The areas of lower contrast will gain a higher contrast without affecting the global characteristic of image.

3.3.4.2 Global Thresholding

The global thresholding is an operation, when a continuous gray scale of an image is reduced into monochrome black & white colors according to the global threshold value. Let 0,1 be a gray scale of such image. If a value of a certain pixel is above the threshold t, the new value of the pixel will be zero. Otherwise, the new value will be one for pixels with values above the threshold t.

Information contained in a bitmap representation of an image is not suitable for processing by computers. Because of this, there is need to describe a character in another way. All instances of the same character should have a similar description. A description of the character is a vector of numeral values, so-called “descriptors”, or “patterns”:

$$\mathbf{x} = (x_0 ,\ldots, x_{n-1})$$

Generally, the description of an image region is based on its *internal* and *external* representation. The internal representation of an image is based on its regional properties, such as color or texture. The external representation is chosen when the primary focus is on shape characteristics. The description of normalized characters is based on its external characteristics because we deal only with properties such as character shape. Then, the vector of descriptors includes characteristics such as number of lines, the amount of horizontal, vertical and diagonal or diagonal edges, and etc. The feature extraction is a process of transformation of data from a bitmap representation into a form of descriptors, which are more suitable for computers.

If we associate similar instances of the same character into the classes, then the descriptors of characters from the same class should be geometrically closed to each other in the vector space. This is a basic assumption for successfulness of the pattern recognition process.

3.3.4.3 Pixel matrix

The simplest way to extract descriptors from a bitmap image is to assign a brightness of each pixel with a corresponding value in the vector of descriptors. Bigger bitmaps produce extremely long vector of descriptors, which is not suitable for recognition. Because of this, size of such processed bitmap is very limited. In addition, this method does not consider geometrical closeness of pixels, as well as its neighboring relations. Two slightly biased instances of the same character in many cases produce very different description vectors.

3.3.4.4 Detection of character edges

The detection of character edges does not consider positioning of each pixel, but only a number of occurrences of individual edge types in a specific region of the character bitmap. Because of this, the resulting vector is invariant towards the intra-regional displacement of the edges, and towards small deformations of characters.

3.3.5 Module 5 – Recognition of characters

CHAPTER 4
EVALUATION AND RESULTS

This chapter explains about the system evaluation and performance of the system.

4.1 EXPERIMENTAL RESULTS

4.1.1 Program execution

The syntax to give input to the recognizer program.

java Main <image_file_name>

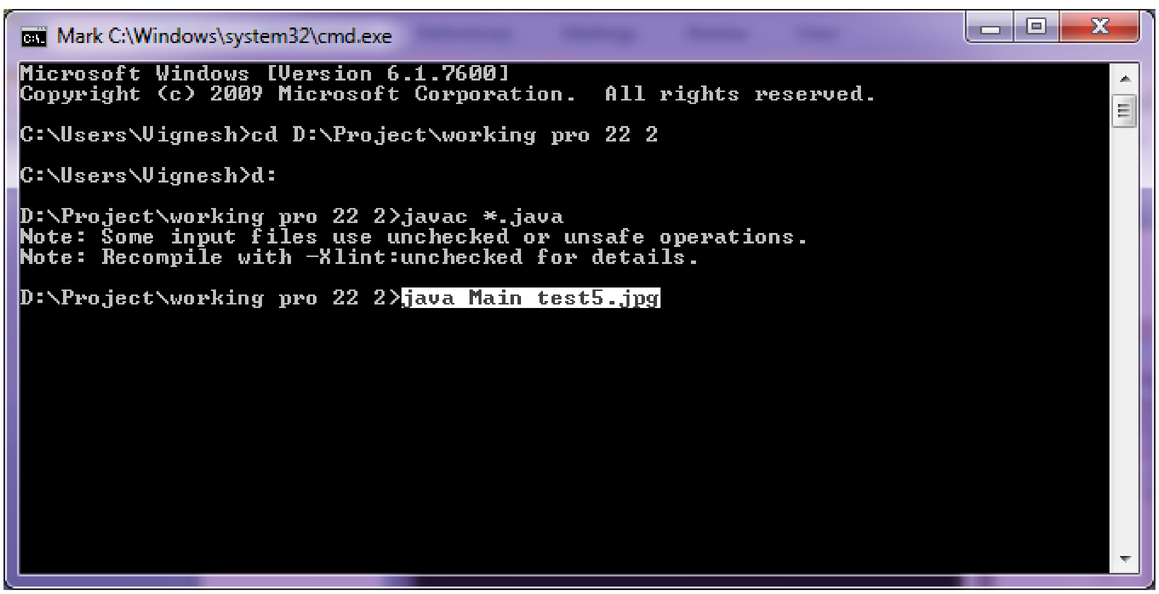


Figure 4.1.1 Execution of Main program to recognize characters in the number plate

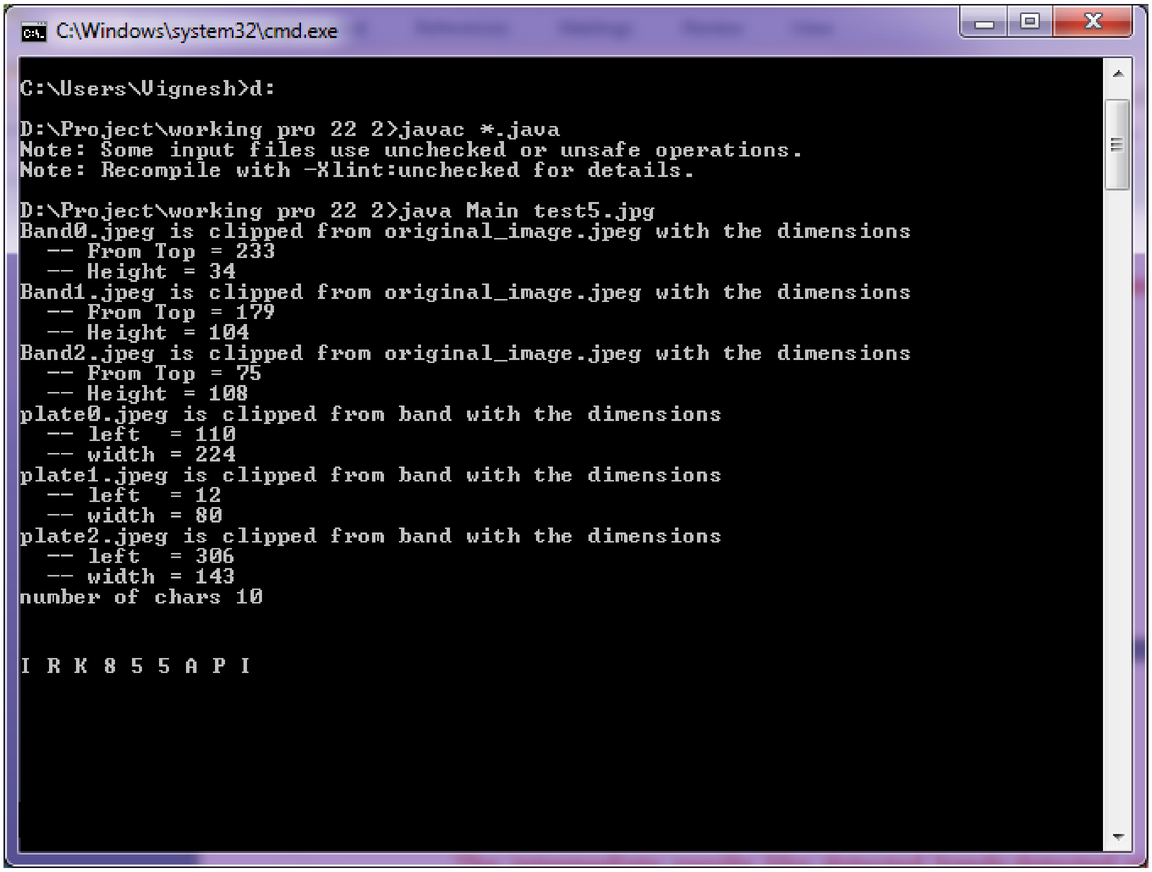


Figure 4.1.2 Recognized characters

4.1.2 Intermediate Results:

The intermediate results like detected bands, detected plates and segmented characters has been stored in the folder “./screenshots”.

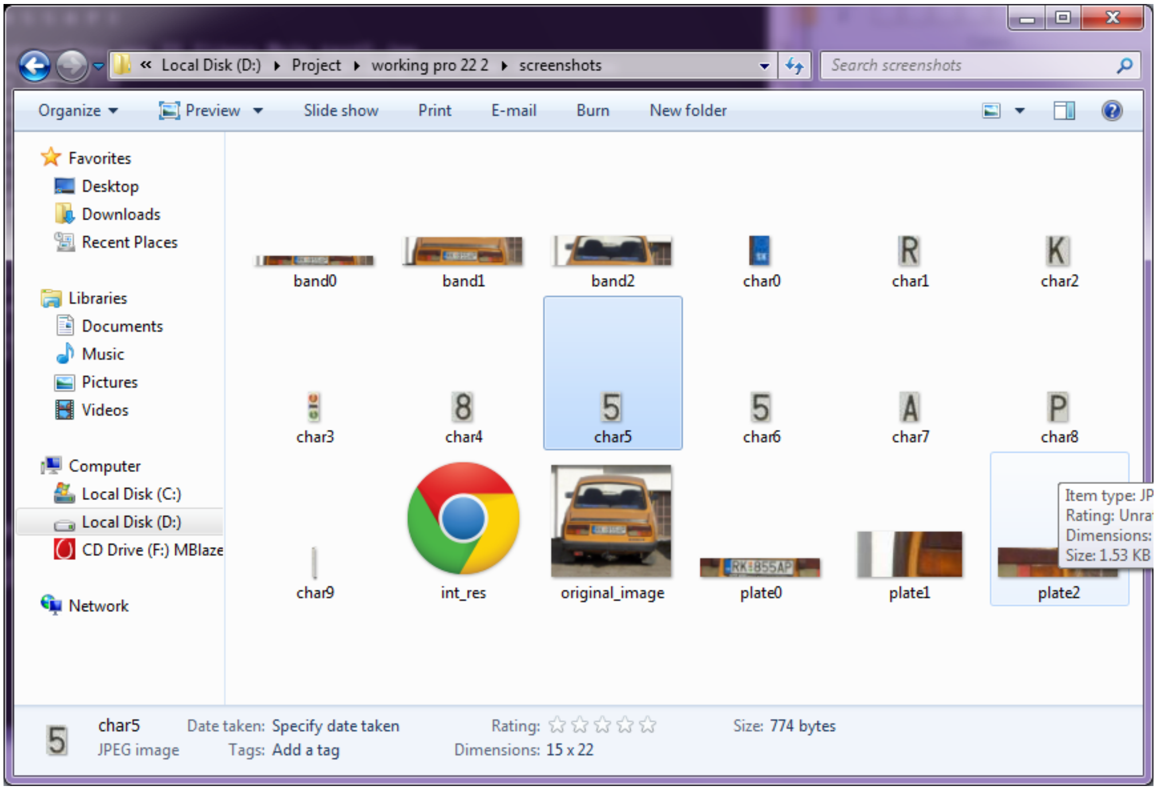


Figure 4.2 Recognized characters

All the intermediate results (images) can be viewed through the HTML page as shown in the figure 4.3

Figure 4.3.1 Original image and detected bands in a HTML page

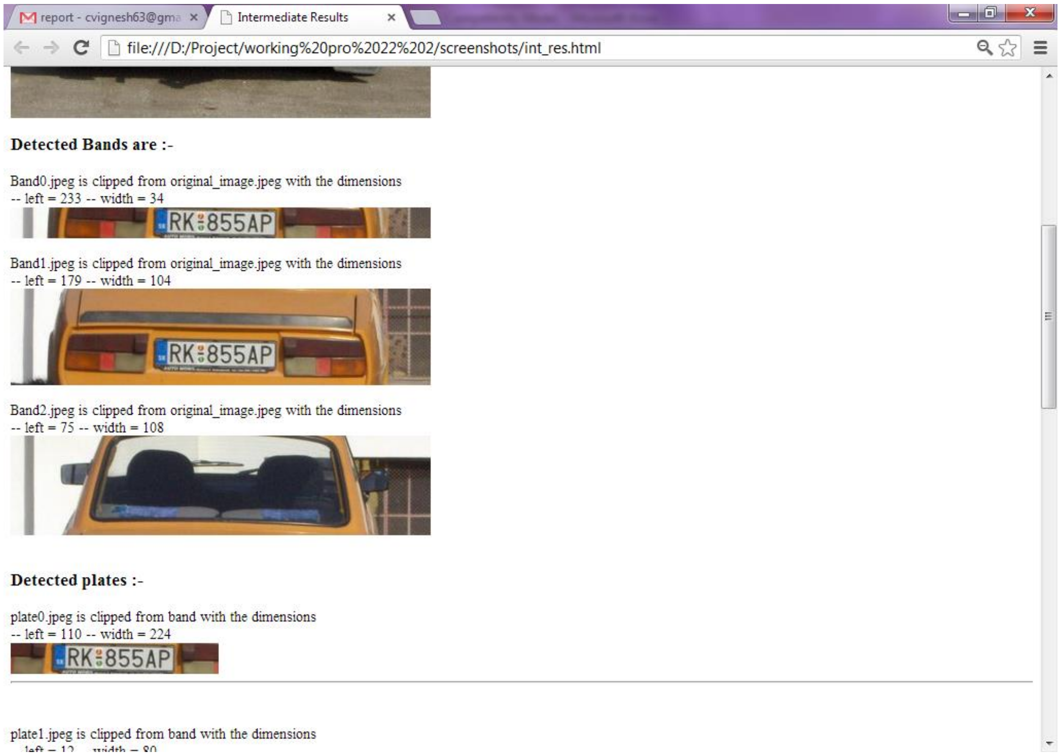


Figure 4.3.2 Detected bands and plates in a HTML page

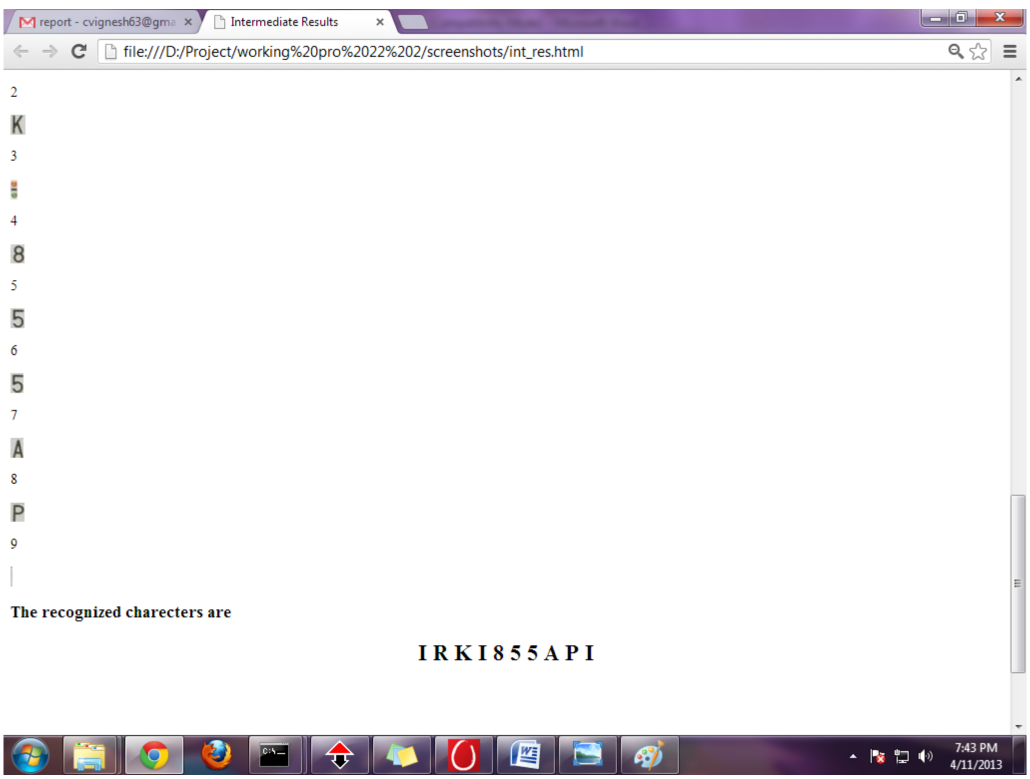


Figure 4.3.3 Segemented characters and recognized characters

4.1.3 Number Plate Detection:

The high density of horizontal and vertical edges on a small area in the input image is considered as a number plate. This process can sometimes detect a wrong area that does not correspond to a number plate. Because of this, we often detect several candidates for the plate by this algorithm, and then we choose the best one by a further analysis. The bands are detected using vertical projection of the input image. The detected bands for the sample image is shown in Fig 4.3.The band clipped image is further clipped using horizontal projection to get the number plates.

Figure 4.4 Detected Bands

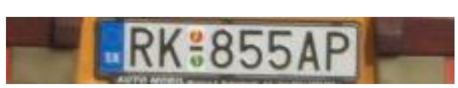


Figure 4.5 Detected plate

4.1.4 Character Segmentation:

We have used a horizontal projection of a detected number plate for the segmentation of the characters. The module also enhances the characters by eliminating undesirable elements such as dots and stretches contains besides the character. Then the extracted characters are used for further processing. The segmented characters are shown in Fig.

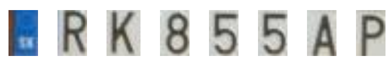


Figure 4.6 Segmented characters

After the segmentation, the file is stored as character file in the same location as the band file as displayed in Figure 4.2.

4.2 PERFORMANCE ANALYSIS

A database consists of different sized JPEG colored images. Totally 100 images are used to test the efficiency of the different algorithms used in this project.

s.no		Correct output count	%
1	Total number of images	100	
2	plates correctly clipped	95	95
3	Characters correctly segmented	83	83
4	detection and elimination of redundant elements within the segmented character (e.g dots or screws)	92	92
5	character recognition		
5.1	K nearest neighbor algorithm	71	71
5.2	Neural networks	77	77

The program works satisfactorily for wide variations in illumination conditions and different types of number plates.

The program correctly identifying (clipping) the plate in the image for most of the cases, even though it is further needed to apply deskewing mechanism.

This system may sometimes detect a wrong area that does not correspond to a number plate. But the system might detect several candidates for the plate to choose the best one by a further analysis.

Character segmentation is getting degraded when it is corrected through deskewing mechanism. However 70 percent of the characters are detected during the character segmentation phase.

Character recognition performance is analyzed for two different algorithms named knn (k-nearest neighbor) algorithm and neural networks that is used to recognize the characters. More number of training sets for neural networks gives a better performance to recognize the characters. If the 70% of the characters are identified correctly by the program, then it is considered as a correct one.

The proposed system is perceived to have drawbacks under situations include poor resolution of the captured images because of the quality of the cameras used or the distance of the vehicle being photographed, lighting or contrast problems because of shadows.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

This chapter summarizes the main components of this system as well as the major conclusions gathered from the results of the evaluation. This section also covers how the system can be further improved.

Thus, a mispronunciation detection speech recognition system has been developed for Tamil language with phoneme as smallest word unit of recognition. A set of Tamil words and their phoneme sequence in stored in a database along with common mispronunciations. If there is any mispronunciation then the appropriate phoneme is replaced in the found phoneme sequence to find the correct word.

In future, the isolated speech recognition may be extended as continuous words speech recognition. Moreover, speaker independent systems can be developed based on the current system. The system can also be upgraded to work in noisy environments.

APPENDIX – 1

The source code, executables, data corpus, tools required and readme file are written onto a CD-ROM and it is attached with this report at the end page.

REFERENCES

[1] Muhammad Tahir Qadri, Muhammad Asif (2009): Automatic Number Plate Recognition System For Vehicle Identification Using Optical Character Recognition, International Conference on Education Technology and Computer.

[2] Chen-Chung Liu, Zhi-Chun Luo(2010): An Extraction Algorithm of Vehicle License Plate Numbers Using Pixel Value Projection and License Plate Calibration, International Symposium on Computer, Communication, Control and Automation.

[3] Peter M. Roth, Martin K¨ostinger, Paul Wohlhart, and Horst Bischof, Josef A. Birchbauer (2010): Automatic Detection and Reading of Dangerous, 2010 Seventh IEEE International Conference on Advanced Video and Signal Based Survillance

[4] Ping Dong, Jie-hui Yang, Jun-jun Dong (2006): The Application and Development Perspective of Number Plate Automatic Recognition Technique.

[5] W. K. I. L. Wanniarachchi, D. U. J. Sonnadara and M. K. Jayananda (2007): License Plate Identification Based on Image Processing Techniques, Second International Conference on Industrial and Information Systems.

[6] Ankush Roy Debarshi Patanjali Ghoshal(2011): Number Plate Recognition for Use in Different Countries Using an Improved Segmentation, IEEE.

[7] B. Raveendran Pillai, Prof. (Dr). Sukesh Kumar. A (2008): A Real-time system for the automatic identification of motorcycle - using Artificial Neural Networks, International Conference on Computing, Communication and Networking.

[8] Pletl Szilveszter, Gálfi Csongor(2010): Parking surveillance and number plate recognition application, IEEE 8th International Symposium on Intelligent Systems and Informatics.

[9] Chirag N. Paunwala, Suprava Patnaik, “A Multiple License Plate Extraction Technique for Complex Background in Indian Traffic Conditions”, In Proceedings of International Journal of Image Processing, vol.4, issue2,2007.

[10] Shishir Kumar, Shashank Agarwal & Kumar Saurabh, “License plate recognition system for Indian vehicles”, IEEE Eighth International Conference on Intelligent transportation System, 2008.

[11] Shidore, Narote, “Number Plate Recognition for Indian Vehicles”,IJCSNS International Journal of Computer Science and Network Security, (2011).

[12] Suri, Walia, Verma, “Vehicle Number Plate Detection using Sobel Edge Detection Technique”, Issue 2, December (2010).

[13] Liu Ying Li Nannan, “Design of License Plate Recognition System Based on the Adaptive Algorithm”, Proceedings of the IEEE International Conference on Automation and Logistics Qingdao, China, 2008.

[14] Roy, A.; Ghoshal, D.P.” Number Plate Recognition for use in different countries using an improved segmentation” IEEE Conference Publications 2011.

[15] “JAVA” URL: http://www.sun.java.com/